

CTCF: insights into insulator function during development

Martin Herold, Marek Bartkuhn and Rainer Renkawitz*

Summary

The genome of higher eukaryotes exhibits a patchwork of inactive and active genes. The nuclear protein CCCTC-binding factor (CTCF) when bound to insulator sequences can prevent undesirable crosstalk between active and inactive genomic regions, and it can also shield particular genes from enhancer function, a role that has many applications in development. Exciting recent work has demonstrated roles for CTCF in, for example, embryonic, neuronal and haematopoietic development. Here, we discuss the underlying mechanisms of developmentally regulated CTCF-dependent transcription in relation to model genes, and highlight genome-wide results indicating that CTCF might play a master role in regulating both activating and repressive transcription events at sites throughout the genome.

Key words: *Igf2/H19*, Cohesin, Homeotic gene, Imprint, Insulator

Introduction

As early as the 1950s, the existence of genomic insulators has been postulated based on several observations, such as the phenomenon of position effect variegation (see Glossary, Box 1), in which gene activity is dependent on its genomic location (Lewis, 1950). These locations were later identified as heterochromatic, or inactive, chromatin regions, which, upon translocation, inversion or deletion of chromosomal fragments, may repress the expression of a neighbouring gene (Baker, 1968). The interpretation of this observation was that the genomic rearrangement might have deleted insulators that normally protect genes from the repressive effect of flanking heterochromatin. To add weight to this notion, the biophysical analysis of the *Drosophila* genome argued for the presence of insulator sequences that separate supercoiled genomic domains (Benyajati and Worcel, 1976). More recently, functional tests with transgenes (see Glossary, Box 1) revealed the existence of three properties of insulators (see Glossary, Box 1): (1) to provide a barrier (or boundary; see Glossary, Box 1) function to prevent repressive heterochromatin from spreading into a neighbouring domain; (2) to provide an enhancer-blocking (see Glossary, Box 1) function when positioned between the enhancer and promoter (see Glossary, Box 1) (Sun and Elgin, 1999), allowing insulators to produce opposite effects either by facilitating the maintenance of a transcriptionally active state or by inhibiting the action of enhancers; (3) to allow three-dimensional looping of genomic regions, a property that is possibly inherent to insulator function, as discussed throughout this review.

The first proteins identified that bind to insulator sequences and mediate the insulating function were boundary element-associated factor of 32 kDa (BEAF32) (Zhao et al., 1995), Su(Hw) (Holdridge and Dorsett, 1991; Geyer and Corces, 1992) and zeste-white 5

(Zw5; dwg – FlyBase) (Gaszner et al., 1999) in *Drosophila*. In vertebrates, CCCTC-binding factor (CTCF) was first shown to mediate insulation (Bell et al., 1999) and was later demonstrated to be highly conserved and also found in *Drosophila* (dCTCF) (Moon et al., 2005). The DNA-binding domains of human and *Drosophila* CTCF are ~60% similar. Interestingly, in *Drosophila*, CTCF is only one of several insulator proteins, whereas in vertebrates, only CTCF is known to mediate insulation function. DNA binding of CTCF is sequence specific and conferred by its zinc finger domain, which in vertebrates and in insects harbours 11 zinc fingers (see Glossary, Box 1) that appear to be differentially used depending on the respective binding site in the genome (Burcin et al., 1997; Boyle et al., 2010).

Here, we provide a comprehensive overview of our current knowledge of the functions of CTCF, shaped by exciting recent developments that are beginning to reveal the molecular mechanisms of CTCF with respect to the chromatin insulation activity required for proper development. Moreover, we highlight the similarities and differences between insulation in *Drosophila* and vertebrates using two well-analysed genomic regions: a homeotic gene locus and the imprinted *Igf2/H19* locus, respectively. Finally, we discuss current information on genome-wide CTCF-binding sites and interpret these in the context of chromatin insulation.

Why do cells need chromatin insulation?

Microscopic inspection of interphase nuclei clearly demonstrated a chromatin ‘patchwork’ in which DNA was packaged as compact heterochromatin (see Glossary, Box 1) or decondensed euchromatin (see Glossary, Box 1). Recent molecular analyses on a genome-wide scale supported the domain model of the linear genome (Fig. 1) (Filion et al., 2010; Ernst et al., 2011). These domains have to be insulated from each other in order to prevent aberrant gene expression, e.g. to prevent the inactive domain from erroneously inhibiting genes in the active domain and vice versa (Fig. 1). Within domains of potentially active genes, the activity of a given gene is controlled by enhancer sequences, which can be found either adjacent to the gene promoter or at a considerable distance either downstream or upstream of the gene. Distance at the level of the linear genome does not impose a problem for enhancer function as the intervening DNA is looped out such that promoter and enhancer will be in close contact (Fig. 1).

The question therefore arises: How does the distant enhancer locate the corresponding promoter and, in particular, how does enhancer ‘homing’ work when two genes with different expression profiles are close together and in a ‘head to head’ orientation (Fig. 1)? Enhancers or locus control regions could be envisaged to recognise and contact their corresponding promoters through a specific composition of interacting proteins bound to the enhancer and to the corresponding promoter. However, in many cases insulators are known to play a role in the three-dimensional folding of chromatin (see below), thereby allowing for or preventing functional contact between enhancer and promoter elements.

Institut für Genetik, Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 58-62, 35392 Giessen, Germany.

*Author for correspondence (rainer.renkawitz@gen.bio.uni-giessen.de)

Box 1. Glossary

Antennapedia complex (ANT-C). A group of five *Drosophila* homeotic genes controlling the development of the head and the anterior part of the thorax.

Barrier. A subgroup of insulator elements found at genomic regions separating an active domain from an inactive domain. Loss of a barrier function often results in inactivation of (part of) the active domain.

Beckwith-Wiedemann syndrome (BWS). A foetal over-growth syndrome that is, in approximately half the cases, caused by a shift in the ratio of *Igf2* and *H19* gene activities (towards increased *Igf2* expression). Such a change is often the result of an imprinting defect.

Bithorax complex (BX-C). A group of three *Drosophila* homeotic genes controlling the development of the abdomen and the posterior part of the thorax.

ChIA-PET (chromatin-interaction analysis with paired-end tag sequencing). A method coupling an initial chromatin-immunoprecipitation with 3C-methodology. Allows the identification of factor-specific long-range chromatin interactions.

ChIP (and ChIP-seq or ChIP-chip). Methods involving chromatin immunoprecipitation (ChIP) and identification of the precipitated DNA by sequencing (ChIP-seq) or by hybridisation to microarrays (ChIP-chip).

Chromosome conformation capture (3C). A method to detect long-distance interactions of chromatin. Involves chromatin crosslinking and subsequent digestion with restriction enzymes and ligation of crosslinked fragments.

Cohesin. A complex of proteins forming a chain structure between paired sister chromatids. Recent studies suggest that chromatin looping between enhancer and promoter regions and between insulators also requires cohesin.

Enhancer. A genomic region that, upon binding of activator proteins, can activate a promoter from a distance. Can be positioned distantly upstream or downstream relative to the promoter.

Enhancer-blocking element. A subgroup of insulator elements that, when bound by an insulator protein, e.g. CTCF, interferes with enhancer-promoter communication.

Euchromatin. Loosely packed chromatin containing active or potentially active genes.

Heterochromatin. Tightly packed chromatin containing genes that are persistently turned off and gene-poor regions enriched with repetitive sequences.

Hox gene cluster. Homeotic (Hox) genes are responsible for developmental patterning along the anteroposterior axis. The linear arrangement of Hox genes on the chromosome reflects their expression along the body axis.

Imprint. An activity status of a gene inherited from one of the parents and maintained in the offspring. A small fraction of human genes, which are highly conserved, is known to be imprinted.

Insulator. An umbrella term for two types of genomic features with either a barrier function or an enhancer-blocking function.

Polycomb response element (PRE). Regulatory DNA elements bound by polycomb group proteins, which are involved in gene repression.

Polytene chromosome. Interphase chromosome consisting of a high number of paired chromatids as found, for example, in salivary glands of insects.

Position effect variegation (PEV). An effect in which a repositioned gene (i.e. one that is placed at a new genomic location) shows variation in gene activity (variegation) from cell to cell.

Promoter. Regulatory region that overlaps with the transcriptional start site of a gene and facilitates transcription. The DNA sequence of the core promoter, the minimal portion of the promoter required to initiate transcription, often contains a string of TATA nucleotides recognised by the TFIID complex.

Silver-Russell syndrome (SRS). Growth disorder causing dwarfism, which is often caused by an increase in *H19* gene activity relative to *Igf2* gene activity.

Transgene. A gene experimentally integrated into the genome, such that this gene is passed on to the offspring.

Zinc finger. A polypeptide motif with cysteine and histidine residues coordinating a zinc ion. Multiples of these motifs are often found in DNA-binding domains.

Sequences involved in insulator function

The first sequences to be functionally tested for insulator activity were chosen because of circumstantial evidence strongly suggesting an insulator function. This was clearly the case for the specialised chromatin structures *scs* and *scs'* (Udvardy et al., 1985), which flank the *Drosophila hsp70* gene and were found to be located at the junctions between decondensed and compact regions on the polytene chromosome (see Glossary, Box 1). Similarly, the *facet-strawberry* sequence is located at a polytene chromosome decondensed interband separating two bands of condensed chromatin (Rykowski et al., 1988). In both cases, functional tests (see Box 2) revealed their insulator activity. Moreover, genetic evidence indicated that some *Drosophila* retrotransposons, such as Gypsy or Idefix, mediate their biological effect by causing insulation at the site of integration, as confirmed by functional tests (Geyer and Corces, 1992; Brasslet et al., 2010). In the case of the two homeotic gene clusters the Antennapedia complex (see Glossary, Box 1) and the Bithorax complex (see Glossary, Box 1), multiple regulatory regions control the segment-specific expression of homeotic genes (see below). Such a complex arrangement of segment-specific regulatory sequences argued for the presence of insulator sequences, which, again, was

confirmed functionally (Hagstrom et al., 1996; Belozerov et al., 2003). Similarly, in vertebrates, the identification of complex expression patterns of adjacent genes or of neighbouring regulatory regions with different activities led to the prediction of insulators at these sites. The first one shown to mediate insulation was the chicken 5'HS4 element, which lies adjacent to the locus control region (LCR) of the β -globin gene cluster and is positioned between a compact chromatin region and the LCR-induced decondensed chromatin (Chung et al., 1993). Finally, an insulator is known to be positioned between the *Igf2* and *H19* genes ensuring their parent of origin-specific expression. This insulator mediates multiple functions related to reading the parental imprint and mediating the imprinted expression of these two genes. Other sequences, for which insulator function has been demonstrated, are listed in Table 1.

A model for insulator function

Given the original observation that the genome is arranged in looped genomic domains (Benyajati and Worcel, 1976) and the postulation that looped domains have to be insulated from each other, the most obvious model for insulator mechanism is that insulators themselves mediate loop formation. As depicted in Fig.

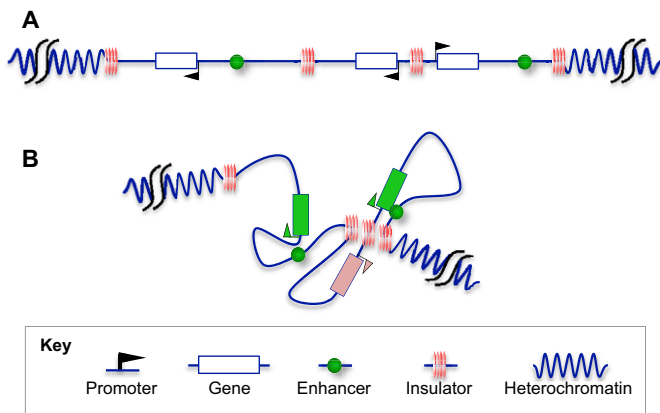


Fig. 1. Schematic overview of the domain model of a linear genome, highlighting insulator locations.

(A) An active chromatin domain is flanked by heterochromatic regions. Insulator positions are indicated at the domain boundaries (where they can mediate border or barrier function of insulators) as well as within the active domain (where they can mediate enhancer-blocking function). It is not known whether both functions are established by similar or different mechanisms. (B) One aspect of insulator function is to organise chromatin looping by promoting contacts between insulators or with other genomic structures. Depending on the linear and three-dimensional arrangement, looping may interfere with enhancer-promoter interactions (thus mediating the enhancer-blocking function of insulators), resulting in an inactive gene (pink gene and promoter), or it may assist in increasing enhancer-promoter contacts, resulting in an active gene (green gene and promoter on right). Gene activation can also be achieved by direct enhancer-promoter interactions (green gene on left) that can occur independently of the presence of an insulator. Insulators are also found between tandem promoters positioned in a head-to-head orientation ensuring that both promoters can be regulated individually.

1, this insulator-mediated looping model does indeed explain how some enhancers are blocked from making contact with promoters, whereas others are ‘pulled’ into the vicinity of a promoter. An alternative model, the decoy model, postulates a competition between insulator and promoter for interaction with the enhancer. In fact, a substantial fraction of CTCF-bound insulators is found to be associated with enhancers (Handoko et al., 2011). However, these looping models cannot account for all cases of enhancer-blocking activity nor can they fully explain how insulator-mediated looped domains affect the barrier mechanisms at play at domain boundaries. Thus, additional features, such as the specific sequence, chromatin context and CTCF co-factors might play a role in insulator function and these are explained below.

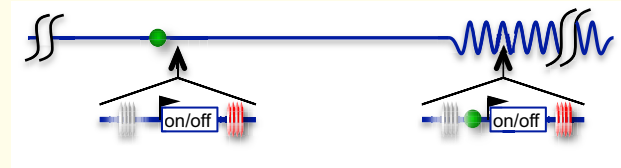
CTCF mediates enhancer blocking at the *Igf2/H19* locus

Imprinting

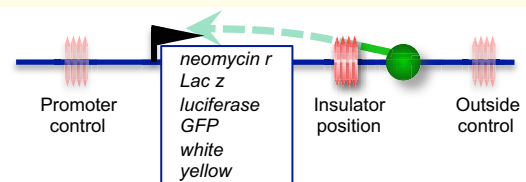
The genetic information of higher organisms, such as insects and vertebrates, is stored in two copies of the genome, the parents contributing one copy each. Usually, both copies (alleles) of a single gene are expressed when the gene is activated. However, in mammalian systems, ~100 genes exhibit a parental imprint (see Glossary, Box 1) resulting in monoallelic expression only. In other words, some genes are only expressed when inherited from the mother and others from the father. Imprinted genes are highly conserved between mouse and human and are often organised in clusters such that paternally and maternally imprinted genes are

Box 2. Techniques to identify insulator sequences

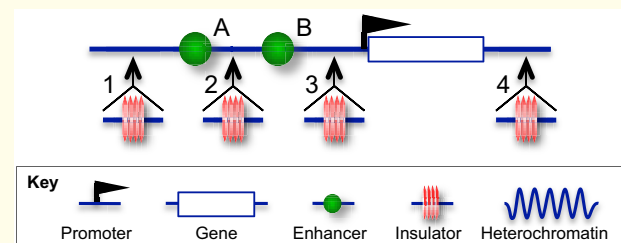
Position effects of reporter genes. Reporter genes not flanked on both sides with insulator sequences show position-specific expression. Enhancer-less reporter genes are expressed when integrated next to an enhancer (left), whereas enhancer-driven reporter genes (right) are repressed when integrated within heterochromatic regions. In cases in which the sequence to be tested (grey insulator) mediates insulation, both position effects caused by genomic enhancers or by heterochromatin are alleviated (Kellum and Schedl, 1991).



Reporter gene tests. A DNA sequence with insulator function should reduce expression of a reporter gene (several alternatives are shown) when the sequence is placed between the enhancer and the gene promoter (insulator position) and not when placed outside of the gene-enhancer unit (outside control) or positioned at the promoter (promoter control).



Genomic gene tests. Insulator sequences interfere with enhancer function when positioned between the gene promoter and the corresponding enhancer of a genomic sequence. Here, two enhancers (A and B) specific for gene expression in tissues A and B, respectively, are shown. Integration of insulators at position 1 or 4 has no effect, at position 2 inhibits gene expression in tissue A only and at position 3 inhibits gene expression in both tissues. Insertion or deletion of insulator sequences can be generated on a large genomic reporter fragment or can be tested by site-directed mutagenesis within the genome.



mixed (for a review, see Ferguson-Smith, 2011). One of the first identified imprinted gene loci was the *insulin-like growth factor 2 (Igf2)/H19* locus (Bartolomei et al., 1991; DeChiara et al., 1991; Ferguson-Smith et al., 1991). Gene knockout studies in mice showed that the *Igf2* gene enhances placental and foetal growth, whereas the *H19* gene, which is expressed as a non-coding RNA, retards foetal growth. Interestingly, both genes in mouse and human are arranged such that the same set of enhancers activates the *H19* gene on the maternal allele and the *Igf2* gene on the paternal allele (Fig. 2A). It was thus proposed that an insulator factor is present at the imprinting control region (ICR) in order to block the enhancer function at the maternal *Igf2* gene (Fig. 2A).

Table 1. List of loci tested for insulator function

Gene	Description	Reference
Drosophila		
ANT-C	The SF1 insulator within the Antennapedia complex.	(Belozerov et al., 2003)
BX-C	<i>Fab-7</i> was the first identified insulator of the Bithorax complex	(Hagstrom et al., 1996; Mihaly et al., 1998)
<i>eve</i>	A binding site for GAGA factor in the <i>eve</i> promoter mediates enhancer blocking	(Ohtsuki and Levine, 1998)
<i>gypsy</i>	Sequences of the <i>gypsy</i> retrotransposon block enhancer function	(Geyer and Corces, 1992; Holdridge et al., 1991)
<i>hsp70</i>	Specialised chromatin structures (<i>scs</i> and <i>scs'</i>) flank the <i>hsp70</i> gene and mediate insulator functions	(Kellum and Schedl, 1991)
<i>Idefix</i>	This retrotransposon has enhancer-blocking activity as well as a barrier function	(Brasset et al., 2010)
<i>Notch</i>	The <i>facet-strawberry</i> sequence insulates the <i>Notch</i> gene from flanking chromatin	(Vazquez and Schedl, 2000)
<i>white</i>	The Wari insulator next to the <i>white</i> gene is present in mini- <i>white</i> constructs	(Chetverina et al., 2008)
–	Five out of eight sites from the genome-wide map of insulators binding CTCF revealed strong enhancer-blocking activity	(Nègre et al., 2011)
Vertebrates		
α -globin	Insulator embedded in the chicken α -globin locus regulates chromatin domain configuration and differential gene expression	(Furlan-Magaril et al., 2010)
α -spectrin	Insulator with barrier-element activity promotes human α -spectrin gene expression in erythroid cells	(Gallagher et al., 2009)
<i>APOB</i>	Human apolipoprotein B flanking sequences serve as boundaries for position-independent transgene expression	(Kalos and Fournier, 1995)
β -globin	5' element of chicken or human β -globin locus exhibits insulator function	(Chung et al., 1993; Li and Stamatoyannopoulos, 1994)
<i>c-Myc (MYC)</i>	The MINE element of the human <i>c-Myc</i> locus mediates barrier function and enhancer blocking	(Gombert et al., 2003)
<i>CFTR</i>	CTCF mediates insulator function at the human <i>CFTR</i> locus	(Blackledge et al., 2007)
<i>DM1</i>	CTCF-binding sites flank CTG/CAG repeats and insulate the human myotonic dystrophy 1 locus	(Filippova et al., 2001)
<i>HLA-DRB1</i> <i>HLA-DQA1</i>	The human major histocompatibility complex class II <i>HLA-DRB1</i> and <i>HLA-DQA1</i> genes are separated by a CTCF-binding enhancer-blocking element	(Majumder et al., 2006)
<i>Igf2</i> and <i>H19</i>	Parent-specific monoallelic expression is in part mediated by an insulator	(Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000)
<i>Jarid1c (Kdm5c)</i>	CTCF site at transition regions between X inactivation and escape functions as insulator	(Filippova et al., 2005)
<i>Kcnq1</i>	A differentially methylated imprinting control region (KvDMR1) within the mouse <i>Kcnq1</i> locus insulates	(Kanduri et al., 2002)
<i>Peg3</i>	Insulator at the paternally expressed mouse <i>Peg3</i> gene is methylated on the maternal allele	(Kim et al., 2003)
rRNA	Insulator within the intergenic spacer of the <i>Xenopus</i> rRNA genes functions as enhancer blocker	(Robinett et al., 1997)
<i>SINE B2</i>	Transcription of a <i>SINE B2</i> repeat as a domain boundary in organogenesis	(Lunyak et al., 2007)
<i>SP10</i>	The testis-specific mouse promoter has insulator function in somatic tissue	(Reddi et al., 2003)
<i>TCRα</i>	A sequence between the T-lymphocyte-specific <i>TCRα</i> locus and the ubiquitously expressed <i>Dad1</i> gene mediates enhancer blocking	(Zhong and Krangel, 1999)
<i>TCRβ</i>	A barrier-type insulator between active and inactive chromatin at the murine <i>TCRβ</i> locus	(Carabana et al., 2011)
–	Constitutively bound CTCF sequences characterised by conservation and enhancer blocking in vertebrates	(Mukhopadhyay et al., 2004; Martin et al., 2011)

Indeed, the insulator factor CTCF was shown by several groups to mediate enhancer blocking at the ICR (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000; Szabo et al., 2000), and this site turned out to be a suitable model for unravelling several insulator features. For example, DNA methylation was found to regulate insulation; the paternal ICR is methylated, whereas the

maternal ICR is not. In this case, only the unmethylated maternal ICR binds CTCF and this in turn blocks *Igf2* gene activation by the enhancer. Deletion of the CTCF-binding sites and maternal transmission results in biallelic expression of *Igf2*, confirming the insulator function of CTCF on the wild-type maternal allele (Engel et al., 2006). Furthermore, Engel and colleagues found that the

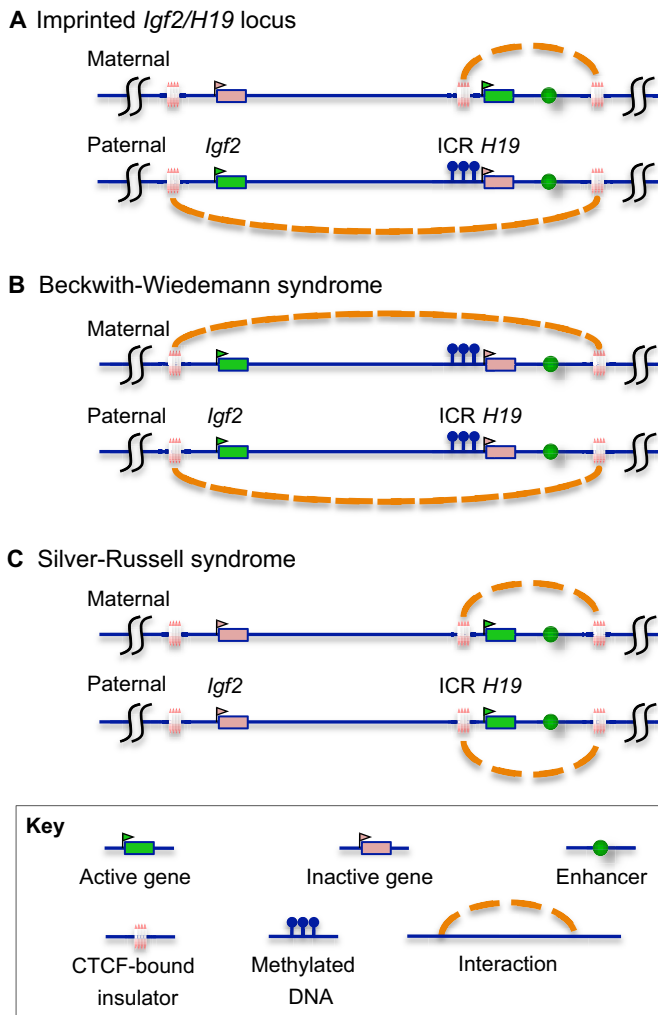


Fig. 2. The imprinted *Igf2/H19* locus as a model for insulator requirement and function. The diagrams depict a selection of the most relevant features within the locus of the *Igf2* and *H19* genes, shown in two dimensions for simplicity. Additional interactions and mechanisms are observed but not shown, as discussed in the text. **(A)** In the normal situation in mouse and human, only a single maternally derived allele shows *H19* gene activity induced by the downstream enhancer. On the paternally derived allele, the DNA of the imprinted control region (ICR) is methylated and prevents CTCF binding such that the insulator function at this site is abrogated. In this situation, the downstream insulator contacts the insulator upstream of the *Igf2* gene and ‘pulls’ the enhancer next to the *Igf2* promoter, which activates the *Igf2* gene. Only some of the long-range interactions are shown. **(B)** In patients with Beckwith-Wiedemann syndrome, DNA methylation at the ICR of both alleles is found as well as an increase in the indicated long-range interaction between insulators, which is associated with biallelic gene activities (*Igf2* expression, *H19* inhibition). **(C)** The situation in patients with Silver-Russell syndrome is characterised by loss of DNA methylation at the ICR. Here, a biallelic pattern is seen with features similar to those of the maternal allele (i.e. *Igf2* inhibition, *H19* expression).

gene locus with a CTCF-binding-site deletion becomes methylated after implantation, suggesting that maintenance of the hypomethylated state requires the presence of the CTCF-binding site. CTCF-binding-site deletion not only demonstrated the enhancer-blocking function of CTCF, but also showed that the

efficient initiation of the adjacent *H19* gene is dependent on CTCF binding. Taken together, these observations suggest that CTCF is also involved in promoter function, a hypothesis that was substantiated by whole-genome analysis (see below). This is a simplified view emphasising insulator function. It is very likely that a concerted action of several mechanisms mediates gene activity at this locus. As an example, direct binding between the maternal ICR and the *Igf2* promoter has also been shown to turn off *Igf2* expression (Li et al., 2008).

The *Igf2/H19* locus in humans

Geneticists have identified imprinted genes in humans based on the parental-specific inheritance of a phenotype caused by the mutant gene, e.g. a mutant *H19* gene is one of several causes of Beckwith-Wiedemann syndrome (BWS, see Glossary, Box 1) only when inherited from the mother. The syndrome is associated with foetal overgrowth and increased risk of tumour formation. However, most cases of BWS are not caused by a gene mutation, but rather by an imprinting defect. For example, one group of patients was found to gain DNA methylation at the ICR (Fig. 2B), resulting in biallelic *Igf2* expression and reduced *H19* activity (Azzi et al., 2009). This BWS phenotype is also observed in cases with microdeletions in the CTCF-binding site of the ICR (Sparago et al., 2004). Thus, loss of CTCF binding at the ICR by either DNA methylation or mutation causes *Igf2* activation and *H19* inactivation. *H19* is not activated because of DNA methylation upstream of the promoter and possibly because of the lack of CTCF binding, which might be required for promoter activation. CTCF sites on both alleles flanking the *Igf2/H19* locus might pull these sites together such that the enhancer can interact with the *Igf2* promoter (Fig. 2B). This scenario was recently confirmed by analysing the three-dimensional arrangement of the CTCF-binding sites of BWS patients in this locus (Nativio et al., 2011). In this study, the downstream CTCF-binding site in control samples was observed to interact with both the ICR and the upstream CTCF site, whereas, in BWS cells, the interaction of this downstream CTCF-binding site with the ICR is reduced, but interaction with the upstream CTCF site is increased (Fig. 2B). Similar to the discussion above of the mouse *Igf2/H19* locus, this explanation and the diagram filter out a single feature that is only part of the molecular mechanism at the human locus, which is likely to involve other interactions and mechanisms, such as DNA methylation, as well.

Further insights into CTCF function in humans came from studies of Silver-Russell syndrome (SRS, see Glossary, Box 1), which is characterised by growth retardation during foetal and postnatal development. In more than 50% of SRS patients, a loss of methylation at the paternal ICR is observed, resulting in biallelic expression of the *H19* gene and reduced expression of *Igf2* (i.e. the converse of the BWS pattern). Analysis of long-distance interaction in SRS cells revealed that the downstream CTCF-binding site exhibits increased interaction with the ICR and reduced contact with the upstream CTCF site (Nativio et al., 2011) (Fig. 2C).

How does CTCF function at the *Igf2/H19* locus?

For the mouse *Igf2/H19* locus, it has been shown that three-dimensional chromatin looping is dependent on CTCF (Kurukuti et al., 2006). An interesting candidate as a mediator of a CTCF-dependent looping function is cohesin (see Glossary, Box 1). Genome-wide analyses of CTCF and cohesin binding resulted in largely overlapping patterns (Parelho et al., 2008; Wendt et al., 2008). Moreover, functional tests revealed that cohesin is bound to CTCF

and is required for chromatin long-distance interaction (Wendt et al., 2008; Hadjur et al., 2009; Nativio et al., 2009). Because cohesin is also involved in sister chromatid cohesion, which is independent of CTCF, it can be envisaged that cohesin confers the linking function and CTCF is responsible for sequence-specific binding. Indeed, it has been demonstrated that the cohesin component SA2 makes direct contact with CTCF, and the ring-forming components of cohesin are recruited by SA2 (Xiao et al., 2011) (Fig. 3A). CTCF complex purification and functional tests showed that the RNA helicase p68 together with its non-coding RNA component SRA are required for insulation at the *Igf2/H19* locus (Yao et al., 2010). These authors proposed that the RNA helicase stabilises cohesin-CTCF interaction and that cohesin mediates chromatin looping.

Thus, an insulation mechanism can be envisaged involving CTCF, p68, SRA-RNA, cohesin and chromatin long-distance interaction (Fig. 3A).

Insulation in *Drosophila*

Several factors, including CTCF, mediate insulator function in *Drosophila*

After identification of the specialised chromatin structures containing insulator sequences in *Drosophila*, namely *scs* and *scs'* (see above), two insulator-binding proteins were identified: zeste-white 5 (Zw5) and boundary element associated factor (BEAF) (Zhao et al., 1995; Gaszner et al., 1999; Roy et al., 2006). In *Drosophila*, additional insulator-binding factors were identified for other insulators. For example, suppressor of hairy wing [Su(Hw)] was found to bind to the well-characterised *gypsy* retrotransposon (Parkhurst et al., 1988), and the GAGA-binding factor (GAF) to the *eve* promoter (Ohtsuki and Levine, 1998). These four proteins are all ubiquitously expressed, harbour a zinc finger domain for direct binding to DNA and are responsible for the insulator activity of specific target sites. Interestingly, the only insulator-binding factor in *Drosophila* with an orthologue in vertebrates is CTCF (Moon et al., 2005) with binding sites in the highly conserved homeotic gene cluster (or Hox cluster, see Glossary, Box 1) (Holohan et al., 2007). The genes within this cluster code for transcription factors that determine body patterning along the anteroposterior axis of an organism. Importantly, the linear arrangement of Hox genes on the chromosome reflects their expression along the body axis, a phenomenon called co-linearity. This arrangement of genes and of segment-specific regulatory sequences argued for the existence of insulator sequences that separate segment-specific regulatory units (Lewis, 1978; Hagstrom et al., 1996; Belozarov et al., 2003). Deletion of these insulator sequences results in a shift of segment identity, such that neighbouring segments show an identical identity (for details, see Maeda and Karch, 2006).

CTCF and the Bithorax complex

In *Drosophila*, one such Hox cluster, known as the Bithorax complex (BX-C, Fig. 4) specifies the third thoracic segment (T3) and all eight abdominal segments (A1-A8) of the fly. BX-C is ~300 kb long and contains three homeotic genes: *Ultrabithorax* (*Ubx*), *abdominal A* (*abd-A*) and *Abdominal B* (*Abd-B*). These three genes are regulated by nine different cis-regulatory domains, each responsible for the identity of the nine segments (Fig. 4) (for a review, see Maeda and Karch, 2006). *Fab-7* was the first insulator identified within BX-C and was found to separate the regulatory domains *iab-6* and *iab-7* from each other (Gyurkovics et al., 1990) and, subsequently, other insulators, such as *Fab-8*, were identified for which enhancer-blocking activity is dependent on CTCF (Moon

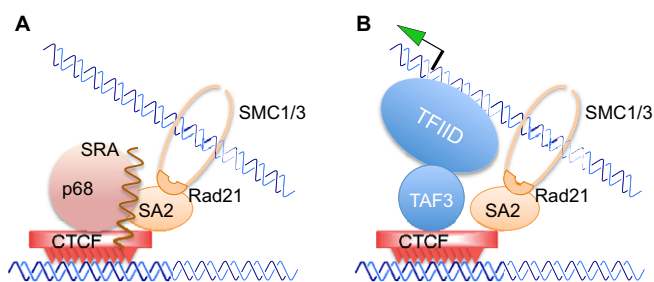


Fig. 3. CTCF and co-factors involved in long-range chromatin interaction. DNA-bound CTCF is depicted together with the cohesin components SA2, Rad21, SMC1 and SMC3, which are known to embrace DNA. The SA2 subunit binds to the C-terminal end of CTCF (Xiao et al., 2011). (A) The CTCF-cohesin complex is stabilised by the RNA helicase p68 and the RNA molecule SRA (red wavy line), which are required for insulation (Yao et al., 2010). (B) CTCF also directly binds TAF3 and mediates DNA looping between distal CTCF-bound sites and TFIIID-bound core promoters, resulting in gene transcription (green arrow) (Liu et al., 2011).

et al., 2005). Furthermore, chromatin immunoprecipitation (ChIP, see Glossary, Box 1) of CTCF revealed binding at most of the boundaries (Fig. 4), thereby underscoring a role for CTCF in dividing the regulatory domains of the BX-C (Holohan et al., 2007; Mohan et al., 2007). Finally, the important role of CTCF for regulatory domain function in the BX-C is underscored by misexpression of *Abd-B* and by homeotic transformations of the abdominal segments in CTCF mutants, resulting in an additional abdominal segment 7 (Gerasimova et al., 2007; Mohan et al., 2007).

A search for co-factors involved in chromatin insulation identified the centrosomal protein 190 (CP190), which binds to Su(Hw) and is required for the enhancer-blocking activity of the *gypsy* retrotransposon (Pai et al., 2004). Genome-wide CP190-binding analysis clearly showed extensive colocalisation of CP190 with Su(Hw), as well as with CTCF. Moreover, all of the CTCF sites in the BX-C are bound by CP190, indicating a role for CP190 in the regulation of the BX-C. In addition to the strong overlap of CP190-binding sites with CTCF and Su(Hw), the insulator-binding factors BEAF and GAF also colocalised with CP190, thus identifying CP190 as a central factor in insulator function (Bartkuhn et al., 2009; Bushey et al., 2009; Nègre et al., 2011). The RNA helicase Rm62 was also found to interact with CP190 in an RNA-dependent manner (Lei and Corces, 2006). Involvement of RNA molecules in chromatin insulation was also suggested from experiments with vertebrate CTCF, in which CTCF interaction with cohesin was found to involve RNA species as well as the RNA helicase p68 (Fig. 3A) (Yao et al., 2010). In both cases, protein-protein interaction with the *Drosophila* Rm62 factor or the similar human p68 factor was dependent on RNA. Whether the RNA component itself is involved in the interaction, or in changing a protein conformation required for other proteins to bind, remains to be shown.

Mechanism of CTCF-mediated insulation within the BX-C

To understand better the underlying mechanisms of insulator function, the three-dimensional structure of the locus was analysed. For this purpose, insulator elements of the BX-C harbouring CTCF-binding sites were cloned into reporter

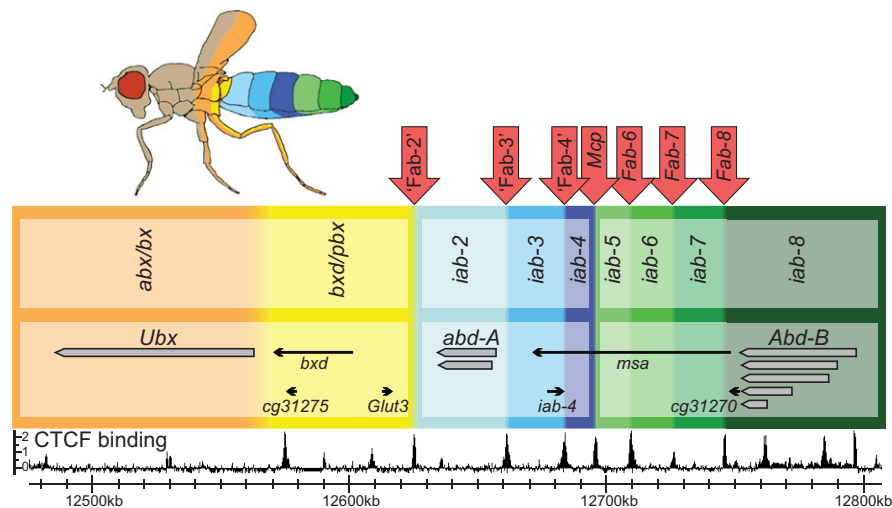


Fig. 4. CTCF binding in the *Drosophila* Bithorax complex. The colour code highlights the nine regulatory domains (*abx/bx*, *bxd/pbx*, *iab-2* to *iab-8*) that are responsible for the segment-specific expression of the three homeotic genes *Ubx*, *abd-A* and *Abd-B* [adapted from Maeda and Karch (Maeda and Karch, 2006)]. The domains shaded in green (*iab-5* to *iab-8*) control the expression of *Abd-B*, the blue domains (*iab-2* to *iab-4*) regulate *abd-A*, and the orange/yellow domains (*abx/bx*, *bxd/pbx*) are responsible for *Ubx* expression within the colour-coded segments of the fly shown in the diagram above. Transcripts of the three homeotic genes are indicated by the grey arrows. Other protein-coding genes (*cg31275*, *Glut3*, *cg31270*) and non-coding genes (*bxd*, *iab-4*, *msa*) are depicted by black arrows. Genetically postulated ('*Fab-2*', '*Fab-3*', '*Fab-4*') or identified (*Mcp*, *Fab-6*, *Fab-7*, *Fab-8*) insulators are marked by red arrows. The CTCF-binding profile is represented as log₂ ratio of CTCF precipitation over input (Bartkuhn, 2009) together with the coordinates on chromosome 3R (bottom). The profile highlights the fact that almost all insulator regions are marked by CTCF binding.

constructs and, after integration in the *Drosophila* genome, were found to interact physically. However, deletion of the CTCF-binding sites only partially reduced the ability of the insulator elements to interact (Kyrchanova et al., 2011). Therefore, other elements or proteins might be involved in the long-range interaction. Indeed, polycomb response elements (PREs, see Glossary, Box 1) were shown to interact *in vivo* in a silenced BX-C background (Lanzuolo et al., 2007), although functional and interaction assays showed that long-range contacts are independent of PREs (Li et al., 2011).

Thus, insulator function in the BX-C locus involves long-distance chromatin interactions although it is difficult to envisage how the enhancers of one domain can be blocked in a specific segment (e.g. *iab-6* in abdominal segment 5), but can be active in another (e.g. *iab-6* in abdominal segment 6).

Therefore, the molecular mechanism regulating the BX-C locus turns out not to be a simple insulation- or relief from insulation-based mechanism, but rather a complex picture is emerging. Complexity is in part caused by additional factors, such as polycomb group proteins, and also by the large number of genomic elements involved in the three-dimensional interaction (for a review, see Bantignies and Cavalli, 2011). Besides insulators, other sequence elements that are involved include enhancers and promoters (see below). Furthermore, promoter-targeting sequences (PTSS) have been shown to allow enhancer elements to bypass the intervening insulators (Zhou and Levine, 1999). Another element, identified in the promoter of *Abd-B*, is the promoter-tethering element (PTE). This element is able to bypass the insulators *Fab-7* and *Fab-8* and to activate *Abd-B* expression via the enhancer element of *iab-5* (Akbari et al., 2007). Therefore, although the insulator function of CTCF and its co-factors at the BX-C locus is well studied, functional details of the underlying mechanisms enabling these factors to confer as well as to block insulation have still to be elucidated.

CTCF in other developmental processes

In addition to mediating developmental gene regulation of the *Igf2/H19* and *Hox* gene loci, CTCF is now known to play a significant role in many other developmental processes (Table 2).

Embryonic development

The importance of CTCF in development is demonstrated by the fact that CTCF-null mice are embryonic lethal (Splinter et al., 2006; Heath et al., 2008). Similarly, CTCF-null *Drosophila* strains do not develop beyond the pupal stage (Gerasimova et al., 2007; Mohan et al., 2007). Furthermore, mouse oocytes depleted of CTCF showed increased methylation of the *H19* ICR, decreased developmental competence, as well as meiotic and mitotic defects in the egg and embryo, respectively. Therefore, CTCF can be considered to be a maternal-effect gene, transcription of which in the oocyte is required for normal embryonic development (Fedoriv et al., 2004; Wan et al., 2008).

Embryonic stem (ES) cells possess the potential to generate all somatic cell types. Differentiation of ES cells serves as a widely used model to analyse specific aspects of embryonic development. Recently, it has been shown that ES cells are highly enriched with the TATA-binding protein associated factor 3 (TAF3) (Liu et al., 2011) and that TAF3 is required for endoderm lineage differentiation. In addition to binding to TFIID at the core promoter (see Glossary, Box 1), TAF3 has been found at a subset of genomic sites marked by CTCF and cohesin. CTCF binds TAF3 directly and mediates DNA looping between distal, CTCF-bound sites and core promoters (Fig. 3B). In this case, CTCF confers gene activation rather than insulation; nevertheless, this example nicely illustrates the idea that CTCF-bound insulators might contact promoters as well (see below).

Embryonic development and ES cell differentiation in mammals requires inactivation of one of the two X-chromosomes in female cells. Major players in this process are two adjacent and head-on

oriented genes *Tsix* and *Xist*. *Xist* upregulation, being essential for the inactive chromosome, is suppressed by *Tsix* activity on the active X-chromosome. CTCF is found to bind to a conserved sequence element positioned between the *Tsix* and *Xist* genes. Depletion of CTCF or deletion of the binding site results in aberrant X inactivation in female cells (Spencer et al., 2011). The interpretation by the authors is that the insulator function ensures the differential activity of these adjacent genes.

As discussed above, homeotic genes define segment identity during embryonic development. As observed in *Drosophila*, loss of CTCF leads to homeotic transformations of the abdominal segments (Mohan et al., 2007). By contrast, in mice, expression of the *HoxD* cluster that is responsible for limb development remains unchanged after conditional *CTCF* gene inactivation. However, the structure of the limbs is impaired owing to increased apoptosis (Soshnikova et al., 2010). This is surprising, as CTCF is bound at sites expected to insulate the different homeotic genes from each other. One interpretation is that establishment of the insulator function occurs early in development, before the conditional *CTCF* deletion was induced.

Neuronal development

Earlier work on the transcriptional regulation of the gene coding for amyloid precursor protein (APP), which is involved in Alzheimer's disease, in rat primary neurons suggests that CTCF is important for the upregulation of APP expression during synaptogenesis (Yang et al., 1999). The CTCF-binding site is located in the promoter, a feature that is found for a substantial number of CTCF-binding sites. Recently, the brain-specific expression of the mouse *protocadherin- α* (*Pcdha*) gene cluster has been analysed (Kehayova et al., 2011). A vast number of *Pcdh* isoforms contributes to the diverse repertoire of neuronal cell-surface identities. A battery of alternative promoters is controlled by a remote enhancer. Both promoters and enhancer are bound by CTCF, which is required for promoter activation (Kehayova et al., 2011). As mono-allelic and combinatorial expression of the alternative promoters is observed, a combination of CTCF-mediated promoter function and insulation could be involved in the selection of promoters in individual neurons.

Haematopoietic differentiation

CTCF is also involved in blood cell differentiation. In maturing human dendritic cells (DCs), CTCF expression is normally reduced, whereas the forced overexpression of CTCF during differentiation of murine, bone marrow-derived DCs leads to reduced proliferation and increased apoptosis (Koesters et al., 2007). During differentiation of B-cells, recombination of the *Igh*

locus occurs, which requires locus contraction and looping. Recent studies concluded that the many CTCF-cohesin sites in the locus are responsible for locus compaction because loss of CTCF was observed to lead to a more loose structure (Degner et al., 2011; Guo et al., 2011a). In addition, *immunoglobulin heavy chain* (*Igh*) recombination is also mediated by CTCF (Guo et al., 2011b). Guo and colleagues identified a key *Igh* recombination regulatory region, which they termed intergenic control region 1. This region harbours two CTCF-binding sites, which, when deleted, result in impaired B-cell development. In particular, these mutations alter the usage of the variable gene segments, induce germline transcription and change the order of rearrangements. The authors propose that these CTCF sites mediate loops with downstream CTCF sites that segregate different heavy chain segments into separate regulatory domains during the heavy chain rearrangement stage of B-cell development. Simultaneously, the enhancer activation of the promoters at the variable segments before rearrangement is blocked (Guo et al., 2011b). Such a dual effect of CTCF in mediating enhancer blocking as well as immunoglobulin rearrangement has been observed for the κ light chain locus also. Here, an alternative strategy has been taken by conditionally deleting the CTCF gene in the B-cell lineage of the mouse (Ribeiro de Almeida et al., 2011). The authors observed that in the absence of CTCF, the *Igk* light chain locus showed a change in the usage of $V\kappa$ gene segments as well as a change in germ line transcription. Chromosome conformation capture (3C, see Glossary, Box 1) experiments demonstrated that CTCF limits interactions of the *Igk* enhancers with the proximal $V\kappa$ gene regions. In this system, no effect was seen on the heavy chain locus, which might be explained by an early epigenetic effect by CTCF, which, upon conditional deletion, does not impair the epigenetic marks set by CTCF previously.

Furthermore, T-cell differentiation and maturation also depend on functional CTCF. At the human locus coding for interferon gamma (*IFNG*), CTCF and cohesin sites form the basis for long-range interactions during differentiation of naive CD4 T cells into T-helper T_H1 cells (Hadjur et al., 2009). Knockout of *Ctcf* in mouse thymocytes leads to cell cycle arrest, thereby suggesting that CTCF acts as a regulator of cell growth and proliferation in $\alpha\beta$ T cells (Heath et al., 2008). In addition, reducing the level of CTCF during differentiation of Th2 cells was found to impair Th2 cytokine expression and T-cell activation (Ribeiro de Almeida et al., 2009). Finally, CTCF was also found to control MHC class II gene expression and to be required for the formation of long-distance chromatin interactions at the MHC class II locus (Majumder et al., 2008).

Table 2. The role of CTCF in development

Tissue or developmental stage	Effect	Reference
CD4+T cells	Cell cycle progression of T cells	(Heath et al., 2008)
Dendritic cells (DC)	DC differentiation and subset development and identity	(Koesters et al., 2007)
<i>Drosophila</i> larvae	Homeotic phenotype	See section on <i>Drosophila</i> CTCF in this Primer and Mohan et al. (Mohan et al., 2007)
Foetal and postnatal development	<i>Igf2/H19</i> imprinting	See section on <i>Igf2/H19</i> in this Primer
Mouse embryo	Imprinting	(Fedoriw et al., 2004; Pant et al., 2004)
Mouse limb development	Apoptosis	(Soshnikova et al., 2010)
Mouse oocyte	Meiosis and mitosis	(Fedoriw et al., 2004; Wan et al., 2008)
Muscle development	CTCF enhances the myogenic potential of MyoD and myogenin	(Delgado-Olguin et al., 2011)
Pro-B cells	Involved in chromatin organisation required for <i>Igh</i> rearrangement	(Degner et al., 2011)
Retina development	Repression of <i>Pax6</i> expression	(Li et al., 2004)

Chromatin flip-flop in a CTCF domain

The Wilms tumor 1 (Wt1) transcription factor regulates mesenchymal-epithelial transition during renal development. The reverse process, epithelial-mesenchymal transition, is essential for generating vascular progenitors in the epicardium. In both cases, Wt1 regulates the *wingless-type MMTV integration site family, member 4* (*Wnt4*) locus. This locus, which is framed by CTCF-binding sites, acquires an active chromatin conformation in kidney cells upon Wt1 action, whereas in epicardial cells Wt1 switches the chromatin architecture into a repressed state (Essafi et al., 2011). The authors call this dichotomous switch a ‘flip-flop’ mechanism. Genes outside the CTCF-marked domain do not change their activities. This suggests that CTCF (and cohesin) provide a boundary function. Additional proof for this function is provided by depletion of CTCF or of cohesin, which results in spreading of the active chromatin architecture in kidney cells into the flanking region, whereas in the case of the epicardium the repressed chromatin extends into the flanking genes (Essafi et al., 2011). This not only highlights the boundary function of CTCF, but also demonstrates that both active and inactive chromatin domains are curbed by CTCF.

Regulation of CTCF function

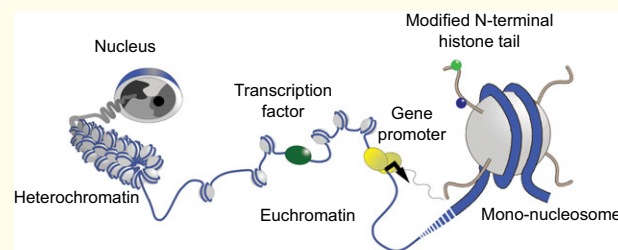
Common to many of these and other developmentally regulated genes is the observation that enhancer activity has to be blocked in some tissues, but conferred in others. Consequently, the enhancer blocking activity of CTCF involved in developmental gene regulation should be regulated. DNA binding of CTCF is one known level of regulation, as demonstrated by DNA methylation or transcription-induced CTCF eviction (Lefevre et al., 2008). Other mechanisms involve changes in CTCF activity without interfering with DNA binding, through synergising DNA-binding factors that have been found to modulate CTCF function (Weth et al., 2010). A recent example of insulator modulation is given by a specific CTCF-binding site in the *Drosophila* genome (Wood et al., 2011). The authors noticed that one of the CTCF co-factors, CP190, was only found at a particular subset of CTCF- or other insulator-binding sites upon treatment with 20-hydroxyecdysone (20-HE), a hormone that binds to the ecdysone receptor complex (ECR-C), which is a classic nuclear receptor complex. At the particular site studied, ECR-C binds upon 20-HE treatment and CP190 is recruited to a CTCF site nearby. This, in addition to a number of constitutive CTCF/CP190 regions, might shield flanking genes from 20-HE action. The authors use 3C-technology to test for long-range chromatin interactions. In fact, chromatin loop formation seems to be slightly increased upon 20-HE addition, indicating a potential increase in insulation.

Finally, post-translational modification of CTCF by phosphorylation (El-Kady and Klenova, 2005) or polyADP-ribosylation (Yu et al., 2004) regulates CTCF function, but precisely how these mechanisms work still remains to be unravelled.

Genome-wide binding studies confirm the proposed functions of CTCF

The aforementioned detailed analyses of CTCF function at model gene loci led to the question of whether similar roles for CTCF can be expected at other genomic locations. Genome-wide detection of CTCF-binding sites is possible using ChIP, to detect interactions of chromatin factors or specifically modified histones with DNA in vivo. In recent years, direct sequencing of the precipitated DNA (via ChIP-Seq, see Glossary, Box 1) in order to obtain genome-

Box 3. Basics of chromatin structure



Eukaryotic DNA is packaged into a histone-DNA complex termed chromatin. In the basic working unit, the nucleosome, DNA is wrapped around an octamer of four core histones H2A, H2B, H3 and H4 in ~1.7 turns, which equals 146 bp of DNA. Neighbouring nucleosomes are separated by a short stretch of DNA of ~20-50 bp termed linker DNA. This configuration can be observed in the most loosely packed ‘beads on a string’ configuration. Packaging the nucleosome core particles into larger aggregates/complexes can further enhance this first level of chromatin compaction. In order to switch between a more densely packed, and therefore repressive, state and the more open, and generally active, state, the cell uses a multitude of functions to modify chromatin such that the required configuration can be established. These functions involve a machinery of factors that modify the histones by covalent modifications of diverse amino acids often located in the N-terminal tails.

A combination of different modification patterns is thought to serve as a binding platform for the recruitment of activating or repressing factors that regulate downstream functions (like transcriptional initiation). A prominent example of such a modification is the triple methylation of lysine 4 of histone H3 (signified as ‘H3K4me3’) often found in the context of active or of poised promoters. Furthermore, acetylation of lysines 9 and 27 of H3 marks active promoters as well as enhancers. By contrast, methylation at lysine 9 (as found in classical heterochromatin) and lysine 27 (polycomb-repressed chromatin) is associated with factors driving chromatin into a repressed state.

wide binding profiles for the factor of interest has become the method of choice (Johnson et al., 2007). CTCF-binding analysis revealed an almost equal distribution of CTCF sites within three classes of binding regions: intragenic, intergenic or in the vicinity of transcriptional start sites (Barski et al., 2007; Guelen et al., 2008; Cuddapah et al., 2009). Genome-wide binding studies to map histone modifications and associated factors for different cell types revealed large blocks of repressed and of active chromatin (Box 3) (Barski et al., 2007; Guelen et al., 2008). Based on their specific composition of associated factors or histone modification marks, these large blocks were further subdivided into specialised chromatin domains (Filion et al., 2010; Roy et al., 2010; Ernst et al., 2011). Combining CTCF binding and chromatin feature analysis revealed that CTCF often demarcates the boundaries of repressed chromatin associated with the nuclear lamina and/or with H3K27me3 adjacent to active regions (Fig. 5) (Barski et al., 2007; Guelen et al., 2008; Cuddapah et al., 2009). These active regions are marked by methylated H3K4 as well as acetylated H2AK5. Importantly, this demarcation of boundaries by CTCF is evolutionarily conserved across species. Despite the fact that there are some general differences in the overall distribution of CTCF profiles in human and *Drosophila*, similarly to the human system, *Drosophila* CTCF has been shown to bind to borders of H3K27me3 domains (Fig. 5). (Bartkuhn et al., 2009; Nègre et al., 2010). Additionally, an evolutionary comparison of the CTCF

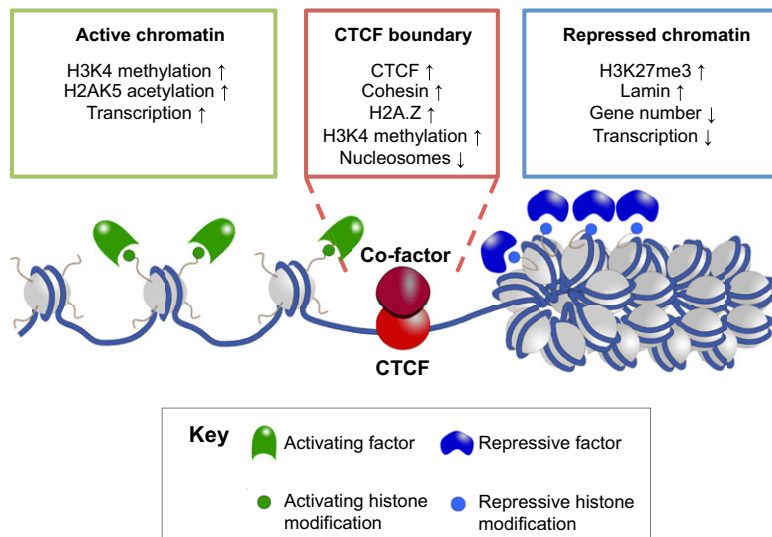


Fig. 5. CTCF boundary elements are associated with specific chromatin features. CTCF-binding sites are characterised by nucleosome depletion. This may be regulated by specific interaction with remodelling factors. Nucleosomes directly associated with CTCF binding have been reported to be marked by generally active marks (like H3K4me3) or exchange histones (like H2A.Z). Furthermore, CTCF boundaries are located in between domains with contrary modification patterns corresponding to contrary activity states. Active chromatin domains are marked by increased acetylation and methylation correlated with increased transcriptional processes. By contrast, repressed chromatin domains are marked by increased H3K27me3 and are low in gene number.

profiles in human, mouse and chicken revealed conserved binding of CTCF to boundaries flanking disease-associated transcription factor genes. This finding suggests a potentially important role for CTCF in diseases associated with the corresponding transcriptional programmes (Martin et al., 2011).

Finally, an additional characteristic of these boundaries is that they are often associated with promoters of actively transcribed genes. A link between promoters and boundary elements or CTCF-binding sites is potentially interesting because these two types of cis-regulatory elements appear to have common features. For example, in *Drosophila*, the CTCF co-factor CP190 is found at a large number of active promoters. Sites co-occupied by CTCF and CP190 are marked by a prominent loss in nucleosome occupancy (Bartkuhn et al., 2009), similar to human genomic binding sites of CTCF (Song et al., 2011). Furthermore, in human cells, CTCF-bound chromatin is associated with the H2A.Z (H2AFZ – Human Gene Nomenclature Database) and H3.3 (H3F3 – Human Gene Nomenclature Database) histone variants, which are known for their association with active promoters (Jin et al., 2009). Finally, there is evidence for CTCF-binding sites being over-represented for mono- and triple-methylated H3K4. Again, active promoters are marked by H3K4me3, whereas H3K4me1 is preferentially associated with enhancers (Fu et al., 2008). These and other findings led to the hypothesis that insulators might have evolved as specialised derivatives of promoters (Raab and Kamakaka, 2010). Indeed, two of the current models for insulator function are in line with this idea. First, promoter-like chromatin structures might function as enhancer decoys, thereby blocking illegitimate enhancer-promoter interactions. Second, promoters might serve as roadblocks to separate repressed from active chromatin. One idea is that histone depletion, as found in promoters, is incompatible with spreading of heterochromatin marks beyond a gap within a row of nucleosomes (Raab and Kamakaka, 2010).

In addition to controlling the biochemical properties of chromatin, it is evident that one of the major CTCF functions is to guide long-range chromatin interactions, which may occur either intra- or inter-chromosomally (Ling et al., 2006; Splinter et al., 2006) and might be stabilised by CTCF-CTCF self-interactions (Pant et al., 2004). Alternatively, CTCF-driven interactions might be enforced by recruitment of the cohesin complex as many studies have demonstrated strongly overlapping binding profiles between CTCF and cohesin

complex members (Parelho et al., 2008; Wendt et al., 2008), and both factors are involved in establishing functional chromatin interactions (Hadjur et al., 2009; Nativio et al., 2009). These and other studies led to speculation that CTCF might be ‘the master weaver of the genome’ (Phillips and Corces, 2009) directing the framework of chromatin interactions in a given cell in order to control numerous downstream processes, such as transcription, replication and repair. This view of CTCF has been strongly supported by a recent study using a genome-wide paired-end tag (PET) approach, ChIA-PET (see Glossary, Box 1). This approach combines an initial chromatin immunoprecipitation with a subsequent 3C-based assay that allows direct sequencing of interactions associated with the immunoprecipitated factor and was recently used to map CTCF-containing chromatin interactions in mouse embryonic stem cells (Handoko et al., 2011). Although far from complete, this method allowed the simultaneous detection of ~1800 long-range chromatin interactions involving CTCF-binding sites. Strikingly, CTCF-containing loops span and separate H3K27me3 repressive domains from H3K4 active domains, indicating that CTCF-binding sites are required for establishing or maintaining such chromatin domains. As described in a previous study (Guelen et al., 2008), many H3K27me3 domains coincide with lamin-associated domains (LADs), suggesting an active role for CTCF in the demarcation of LADs from active chromatin. Furthermore, this study revealed an interesting new aspect of CTCF function, as many of the identified loops connect a CTCF-binding site with an enhancer, where enhancers are not usually bound by CTCF directly. This finding suggests that CTCF not only blocks enhancer-promoter communications but, by contrast, might play the opposite role, too. As some of the identified interactions between enhancers and CTCF target sites involved CTCF target sites located next to promoters, it is likely that CTCF might function to tether remote enhancers to their cognate target genes.

The question then arises of how much influence CTCF has on determining the higher-order chromatin structure in the cell nucleus. To date, the answer is still not completely clear. However, although it is likely that other factors also contribute to this structural process, the finding that a large number of CTCF-associated interactions are correlated with the underlying chromatin domain landscape as well as with coordinated gene expression is a strong argument to posit CTCF as a ‘genome organiser’.

Conclusions

Throughout the past two decades of CTCF research, a developmental role for CTCF was postulated and was evident from early on. This role has been clearly confirmed by *CTCF* gene deletion in mice and *Drosophila*, which results in developmental arrest. Today's genome-wide mapping of CTCF-binding sites, chromatin features and three-dimensional chromatin contacts support the view that CTCF is a 'master' regulator involved in many key developmental processes at both the functional and the molecular level. Currently, it remains to be shown whether the promoter-like features of CTCF-binding sites, as well as its chromatin-looping ability, are related to the barrier function of an insulator and/or to enhancer-blocking activity.

Another open question is that of the regulation of CTCF-mediated insulation. For example, as exemplified with the Bithorax locus, insulation of the *Abd-B* gene from segment-specific enhancer elements has to be overcome in the particular segment requiring this enhancer activity. Furthermore, there are many such situations, mostly in development, when insulator activity mediated by CTCF has to be regulated. Besides DNA binding, the ability of CTCF to interact with other factors is likely to be regulated by modification of CTCF. The mediators and mechanisms that govern these modes of regulation, however, remain to be revealed.

Acknowledgements

We thank Reinhard Dammann for critically reading the manuscript and for valuable suggestions.

Funding

Research in our laboratory was supported by the German Research Foundation (DFG).

Competing interests statement

The authors declare no competing financial interests.

References

- Akbari, O. S., Schiller, B. J., Goetz, S. E., Ho, M. C., Bae, E. and Drewell, R. A. (2007). The abdominal-B promoter tethering element mediates promoter-enhancer specificity at the *Drosophila* bithorax complex. *Fly (Austin)* **1**, 337-339.
- Azzi, S., Rossignol, S., Steunou, V., Sas, T., Thibaud, N., Danton, F., Le Jule, M., Heinrichs, C., Cabrol, S., Gicquel, C. et al. (2009). Multilocus methylation analysis in a large cohort of 11p15-related foetal growth disorders (Russell Silver and Beckwith Wiedemann syndromes) reveals simultaneous loss of methylation at paternal and maternal imprinted loci. *Hum. Mol. Genet.* **18**, 4724-4733.
- Baker, W. K. (1968). Position-effect variegation. *Adv. Genet.* **14**, 133-169.
- Bantignies, F. and Cavalli, G. (2011). Polycomb group proteins: repression in 3D. *Trends Genet.* **27**, 454-464.
- Barski, A., Cuddapah, S., Cui, K., Roh, T., Schones, D., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837.
- Bartkuhn, M., Straub, T., Herold, M., Herrmann, M., Rathke, C., Saumweber, H., Gilfillan, G. D., Becker, P. B. and Renkawitz, R. (2009). Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J.* **28**, 877-888.
- Bartolomei, M. S., Zemel, S. and Tilghman, S. M. (1991). Parental imprinting of the mouse H19 gene. *Nature* **351**, 153-155.
- Bell, A. C. and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**, 482-485.
- Bell, A. C., West, A. G. and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387-396.
- Belozero, V. E., Majumder, P., Shen, P. and Cai, H. N. (2003). A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of *Drosophila*. *EMBO J.* **22**, 3113-3121.
- Benyajati, C. and Worcel, A. (1976). Isolation, characterization, and structure of the folded interphase genome of *Drosophila melanogaster*. *Cell* **9**, 393-407.
- Blackledge, N. P., Carter, E. J., Evans, J. R., Lawson, V., Rowntree, R. K. and Harris, A. (2007). CTCF mediates insulator function at the CFTR locus. *Biochem. J.* **408**, 267-275.
- Boyle, A. P., Song, L., Lee, B.-K., London, D., Keefe, D., Birney, E., Iyer, V. R., Crawford, G. E. and Furey, T. S. (2010). High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. *Genome Res.* **21**, 456-464.
- Brasset, E., Hermant, C., Jensen, S. and Vaury, C. (2010). The Idefix enhancer-blocking insulator also harbors barrier activity. *Gene* **450**, 25-31.
- Burcin, M., Arnold, R., Lutz, M., Kaiser, B., Runge, D., Lottspeich, F., Filippova, G. N., Lobanenko, V. V. and Renkawitz, R. (1997). Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF. *Mol. Cell. Biol.* **17**, 1281-1288.
- Bushey, A. M., Ramos, E. and Corces, V. G. (2009). Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev.* **23**, 1338-1350.
- Carabana, J., Watanabe, A., Hao, B. and Krangel, M. S. (2011). A barrier-type insulator forms a boundary between active and inactive chromatin at the murine TCRbeta locus. *J. Immunol.* **186**, 3556-3562.
- Chetverina, D., Savitskaya, E., Maksimenko, O., Melnikova, L., Zaytseva, O., Parshikov, A., Galkin, A. V. and Georgiev, P. (2008). Red flag on the white reporter: a versatile insulator abuts the white gene in *Drosophila* and is omnipresent in mini-white constructs. *Nucleic Acids Res.* **36**, 929-937.
- Chung, J. H., Whiteley, M. and Felsenfeld, G. (1993). A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74**, 505-514.
- Cuddapah, S., Jothi, R., Schones, D. E., Roh, T. Y., Cui, K. and Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* **19**, 24-32.
- DeChiara, T. M., Robertson, E. J. and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64**, 849-859.
- Degner, S. C., Verma-Gaur, J., Wong, T. P., Bossen, C., Iverson, G. M., Torkamani, A., Vettermann, C., Lin, Y. C., Ju, Z., Schulz, D. et al. (2011). CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc. Natl. Acad. Sci. USA* **108**, 9556-9571.
- Delgado-Olguin, P., Brand-Arzamendi, K., Scott, I. C., Jungblut, B., Stainier, D. Y., Bruneau, B. G. and Recillas-Targa, F. (2011). CTCF promotes muscle differentiation by modulating the activity of myogenic regulatory factors. *J. Biol. Chem.* **286**, 12483-12494.
- El-Kady, A. and Klenova, E. (2005). Regulation of the transcription factor, CTCF, by phosphorylation with protein kinase CK2. *FEBS Lett.* **579**, 1424-1434.
- Engel, N., Thorvaldsen, J. L. and Bartolomei, M. S. (2006). CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus. *Hum. Mol. Genet.* **15**, 2945-2954.
- Ernst, J., Kheradpour, P., Mikkelson, T. S., Shores, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R., Coyne, M. et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43-49.
- Essafi, A., Webb, A., Berry, R. L., Slight, J., Burn, S. F., Spraggon, L., Velecela, V., Martinez-Estrada, O. M., Wiltshire, J. H., Roberts, S. G. et al. (2011). A wt1-controlled chromatin switching mechanism underpins tissue-specific wnt4 activation and repression. *Dev. Cell* **21**, 559-574.
- Fedorow, A. M., Stein, P., Svoboda, P., Schultz, R. M. and Bartolomei, M. S. (2004). Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science* **303**, 238-240.
- Ferguson-Smith, A. C. (2011). Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **12**, 565-575.
- Ferguson-Smith, A. C., Cattanach, B. M., Barton, S. C., Beechey, C. V. and Surani, M. A. (1991). Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature* **351**, 667-670.
- Filion, G. J., van Bommel, J. G., Braunschweig, U., Talhout, W., Kind, J., Ward, L. D., Brugman, W., de Castro, I. J., Kerkhoven, R. M., Bussemaker, H. J. et al. (2010). Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143**, 212-224.
- Filippova, G. N., Thienes, C. P., Penn, B. H., Cho, D. H., Hu, Y. J., Moore, J. M., Klesert, T. R., Lobanenko, V. V. and Tapscott, S. J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.* **28**, 335-343.
- Filippova, G., Cheng, M., Moore, J., Truong, J., Hu, Y., Tsuchiya, K. and Disteche, C. (2005). Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Dev. Cell* **8**, 31-42.
- Fu, Y., Sinha, M., Peterson, C. L. and Weng, Z. (2008). The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.* **4**, e1000138.
- Furlan-Magaril, M., Rebollar, E., Guerrero, G., Fernández, A., Moltó, E., González-Buendía, E., Cantero, M., Montoliu, L. and Recillas-Targa, F. (2010). An insulator embedded in the chicken [alpha]-globin locus regulates chromatin domain configuration and differential gene expression. *Nucleic Acids Res.* **39**, 89-103.
- Gallagher, P. G., Nilson, D. G., Steiner, L. A., Maksimova, Y. D., Lin, J. Y. and Bodine, D. M. (2009). An insulator with barrier-element activity promotes alpha-spectrin gene expression in erythroid cells. *Blood* **113**, 1547-1554.

- Gaszner, M., Vazquez, J. and Schedl, P. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. *Genes Dev.* **13**, 2098-2107.
- Gerasimova, T. I., Lei, E. P., Bushey, A. M. and Corces, V. G. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in *Drosophila*. *Mol. Cell* **28**, 761-772.
- Geyer, P. K. and Corces, V. G. (1992). DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev.* **6**, 1865-1873.
- Gombert, W. M., Farris, S. D., Rubio, E. D., Morey-Rosler, K. M., Schubach, W. H. and Krumm, A. (2003). The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain. *Mol. Cell. Biol.* **23**, 9338-9348.
- Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W. et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948-951.
- Guo, C., Gerasimova, T., Hao, H., Ivanova, I., Chakraborty, T., Selimyan, R., Oltz, E. M. and Sen, R. (2011a). Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell* **147**, 332-343.
- Guo, C., Yoon, H. S., Franklin, A., Jain, S., Ebert, A., Cheng, H. L., Hansen, E., Despo, O., Bossen, C., Vettermann, C. et al. (2011b). CTCF-binding elements mediate control of V(D)J recombination. *Nature* **477**, 424-430.
- Gyurkovics, H., Gausz, J., Kummer, J. and Karch, F. (1990). A new homeotic mutation in the *Drosophila* bithorax complex removes a boundary separating two domains of regulation. *EMBO J.* **9**, 2579-2585.
- Hadjur, S., Williams, L. M., Ryan, N. K., Cobb, B. S., Sexton, T., Fraser, P., Fisher, A. G. and Merkenschlager, M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* **460**, 410-413.
- Hagstrom, K., Muller, M. and Schedl, P. (1996). Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the *Drosophila* bithorax complex. *Genes Dev.* **10**, 3202-3215.
- Handoko, L., Xu, H., Li, G., Ngan, C. Y., Chew, E., Schnapp, M., Lee, C. W., Ye, C., Ping, J. L., Mulawadi, F. et al. (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* **43**, 630-638.
- Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Lovorse, J. M. and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**, 486-489.
- Heath, H., De Almeida, C. R., Sleutels, F., Dingjan, G., Van De Nobelen, S., Jonkers, I., Ling, K.-W., Gribnau, J., Renkawitz, R., Grosveld, F. et al. (2008). CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J.* **27**, 2839-2850.
- Holdridge, C. and Dorsett, D. (1991). Repression of hsp70 heat shock gene transcription by the suppressor of hairy-wing protein of *Drosophila* melanogaster. *Mol. Cell. Biol.* **11**, 1894-1900.
- Holahan, E. E., Kwong, C., Adryan, B., Bartkuhn, M., Herold, M., Renkawitz, R., Russell, S. and White, R. (2007). CTCF Genomic Binding Sites in *Drosophila* and the Organisation of the Bithorax Complex. *PLoS Genet.* **3**, e112.
- Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K. and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.* **41**, 941-945.
- Johnson, D. S., Mortazavi, A., Myers, R. M. and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497-1502.
- Kalos, M. and Fournier, R. E. (1995). Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain. *Mol. Cell. Biol.* **15**, 198-207.
- Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C. F., Wolffe, A., Ohlsson, R. and Lobanenko, V. V. (2000). Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.* **10**, 853-856.
- Kanduri, C., Fitzpatrick, G., Mukhopadhyay, R., Kanduri, M., Lobanenko, V., Higgins, M. and Ohlsson, R. (2002). A differentially methylated imprinting control region within the Kcnq1 locus harbors a methylation-sensitive chromatin insulator. *J. Biol. Chem.* **277**, 18106-18110.
- Kehayova, P., Monahan, K., Chen, W. and Maniatis, T. (2011). Regulatory elements required for the activation and repression of the protocadherin- α gene cluster. *Proc. Natl. Acad. Sci. USA* **108**, 17195-17200.
- Kellum, R. and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**, 941-950.
- Kim, J., Kollhoff, A., Bergmann, A. and Stubbs, L. (2003). Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum. Mol. Genet.* **12**, 233-245.
- Koesters, C., Unger, B., Bilic, I., Schmidt, U., Bluml, S., Lichtenberger, B., Schreiber, M., Stockl, J. and Elmeier, W. (2007). Regulation of dendritic cell differentiation and subset distribution by the zinc finger protein CTCF. *Immunol. Lett.* **109**, 165-174.
- Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenko, V., Reik, W. and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl. Acad. Sci. USA* **103**, 10684-10689.
- Kyrchanova, O., Ivlieva, T., Toshchakov, S., Parshikov, A., Maksimenko, O. and Georgiev, P. (2011). Selective interactions of boundaries with upstream region of Abd-B promoter in *Drosophila* bithorax complex and role of dCTCF in this process. *Nucleic Acids Res.* **39**, 3042-3052.
- Lanzuolo, C., Roue, V., Dekker, J., Bantignies, F. and Orlando, V. (2007). Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat. Cell. Biol.* **9**, 1167-1174.
- Lefevre, P., Witham, J., Lacroix, C. E., Cockerill, P. N. and Bonifer, C. (2008). The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Mol. Cell* **32**, 129-139.
- Lei, E. P. and Corces, V. G. (2006). RNA interference machinery influences the nuclear organization of a chromatin insulator. *Nat. Genet.* **38**, 936-941.
- Lewis, E. B. (1950). The phenomenon of position effect. *Adv. Genet.* **3**, 73-115.
- Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-570.
- Li, H.-B., Muller, M., Bahechar, I. A., Kyrchanova, O., Ohno, K., Georgiev, P. and Pirrotta, V. (2011). Insulators, not polycomb response elements, are required for long-range interactions between polycomb targets in *Drosophila* *Melanogaster*. *Mol. Cell. Biol.* **31**, 616-625.
- Li, Q. and Stamatoyannopoulos, G. (1994). Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. *Blood* **84**, 1399-1401.
- Li, T., Lu, Z. and Lu, L. (2004). Regulation of eye development by transcription control of CCCTC binding factor (CTCF). *J. Biol. Chem.* **279**, 27575-27583.
- Li, T., Hu, J. F., Qiu, X., Ling, J., Chen, H., Wang, S., Hou, A., Vu, T. H. and Hoffman, A. R. (2008). CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. *Mol. Cell. Biol.* **28**, 6473-6482.
- Ling, J. Q., Li, T., Hu, J. F., Vu, T. H., Chen, H. L., Qiu, X. W., Cherry, A. M. and Hoffman, A. R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* **312**, 269-272.
- Liu, Z., Scannell, D. R., Eisen, M. B. and Tjian, R. (2011). Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* **146**, 720-731.
- Lunyak, V. V., Prefontaine, G. G., Nunez, E., Cramer, T., Ju, B. G., Ohgi, K. A., Hutt, K., Roy, R., Garcia-Diaz, A., Zhu, X. et al. (2007). Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* **317**, 248-251.
- Maeda, R. K. and Karch, F. (2006). The ABC of the BX-C: the bithorax complex explained. *Development* **133**, 1413-1422.
- Majumder, P., Gomez, J. A. and Boss, J. M. (2006). The human major histocompatibility complex class II HLA-DRB1 and HLA-DQA1 genes are separated by a CTCF-binding enhancer-blocking element. *J. Biol. Chem.* **281**, 18435-18443.
- Majumder, P., Gomez, J. A., Chadwick, B. P. and Boss, J. M. (2008). The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. *J. Exp. Med.* **205**, 785-798.
- Martin, D., Pantoja, C., Miñán, A. F., Valdes-Quezada, C., Moltó, E., Matesanz, F., Bogdanovi, O., de la Calle-Mustienes, E., Dominguez, O., Taher, L. et al. (2011). Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nat. Struct. Mol. Biol.* **18**, 708-714.
- Mihaly, J., Hogga, I., Barges, S., Galloni, M., Mishra, R. K., Hagstrom, K., Muller, M., Schedl, P., Sipos, L., Gausz, J. et al. (1998). Chromatin domain boundaries in the Bithorax complex. *Cell Mol. Life Sci.* **54**, 60-70.
- Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heintz, N., Bardenhagen, I., Leers, J., White, R. A. H., Renkawitz-Pohl, R., Saumweber, H. et al. (2007). The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J.* **26**, 4203-4214.
- Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., Smith, S. T., Munhall, A., Grewe, B., Bartkuhn, M., Arnold, R. et al. (2005). CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep.* **6**, 165-170.
- Mukhopadhyay, R., Yu, W., Whitehead, J., Xu, J., Lezcano, M., Pack, S., Kanduri, C., Kanduri, M., Ginjala, V., Vostrov, A. et al. (2004). The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res.* **14**, 1594-1602.
- Nativio, R., Wendt, K. S., Ito, Y., Huddleston, J. E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J. M. and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet.* **5**, e1000739.
- Nativio, R., Sparago, A., Ito, Y., Weksberg, R., Riccio, A. and Murrell, A. (2011). Disruption of genomic neighbourhood at the imprinted IGF2-H19 locus in Beckwith-Wiedemann syndrome and Silver-Russell syndrome. *Hum. Mol. Genet.* **20**, 1363-1374.
- Nègre, N., Brown, C. D., Shah, P. K., Kheradpour, P., Morrison, C. A., Henikoff, J. G., Feng, X., Ahmad, K., Russell, S., White, R. A. et al. (2010). A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet.* **6**, e1000814.

- Nègre, N., Brown, C. D., Ma, L., Bristow, C. A., Miller, S. W., Wagner, U., Kheradpour, P., Eaton, M. L., Loriaux, P., Sealfon, R. et al. (2011). A cis-regulatory map of the Drosophila genome. *Nature* **471**, 527-531.
- Ohtsuki, S. and Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo. *Genes Dev.* **12**, 3325-3330.
- Pai, C. Y., Lei, E. P., Ghosh, D. and Corces, V. G. (2004). The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol. Cell* **16**, 737-748.
- Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., Klenova, E., Lobanekov, V. and Ohlsson, R. (2004). Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Mol. Cell. Biol.* **24**, 3497-3504.
- Parelho, V., Hadjir, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T. et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422-433.
- Parkhurst, S. M., Harrison, D. A., Remington, M. P., Spana, C., Kelley, R. L., Coyne, R. S. and Corces, V. G. (1988). The Drosophila su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. *Genes Dev.* **2**, 1205-1215.
- Phillips, J. E. and Corces, V. G. (2009). CTCF: master weaver of the genome. *Cell* **137**, 1194-1211.
- Raab, J. R. and Kamakaka, R. T. (2010). Insulators and promoters: closer than we think. *Nat. Rev. Genet.* **11**, 439-446.
- Reddi, P. P., Shore, A. N., Shapiro, J. A., Anderson, A., Stoler, M. H. and Acharya, K. K. (2003). Spermatid-specific promoter of the SP-10 gene functions as an insulator in somatic cells. *Dev. Biol.* **262**, 173-182.
- Ribeiro de Almeida, C., Heath, H., Krpic, S., Dingjan, G. M., van Hamburg, J. P., Bergen, I., van de Nobelen, S., Sleutels, F., Grosveld, F., Galjart, N. et al. (2009). Critical role for the transcription regulator CCCTC-binding factor in the control of Th2 cytokine expression. *J. Immunol.* **182**, 999-1010.
- Ribeiro de Almeida, C., Stadhouders, R., de Bruijn, M. J., Bergen, I. M., Thongjuea, S., Lenhard, B., van Ijcken, W., Grosveld, F., Galjart, N., Soler, E. et al. (2011). The DNA-Binding protein CTCF limits proximal V κ recombination and restricts kappa enhancer interactions to the immunoglobulin kappa light chain locus. *Immunity* **35**, 501-513.
- Robinett, C. C., O'Connor, A. and Dunaway, M. (1997). The repeat organizer, a specialized insulator element within the intergenic spacer of the Xenopus rRNA genes. *Mol. Cell. Biol.* **17**, 2866-2875.
- Roy, S., Tan, Y. Y. and Hart, C. M. (2006). A genetic screen supports a broad role for the Drosophila insulator proteins BEAF-32A and BEAF-32B in maintaining patterns of gene expression. *Mol. Genet. Genomics* **277**, 273-286.
- Roy, S., Ernst, J., Kharchenko, P. V., Kheradpour, P., Negre, N., Eaton, M. L., Landolin, J. M., Bristow, C. A., Ma, L., Lin, M. F. et al. (2010). Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* **330**, 1787-1797.
- Rykowski, M. C., Parmelee, S. J., Agard, D. A. and Sedat, J. W. (1988). Precise determination of the molecular limits of a polytene chromosome band: regulatory sequences for the Notch gene are in the interband. *Cell* **54**, 461-472.
- Song, L., Zhang, Z., Grasdeder, L. L., Boyle, A. P., Giresi, P. G., Lee, B. K., Sheffield, N. C., Graf, S., Huss, M., Keefe, D. et al. (2011). Open chromatin defined by DNase I and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res.* **21**, 1757-1767.
- Soshnikova, N., Montavon, T., Leleu, M., Galjart, N. and Duboule, D. (2010). Functional analysis of CTCF during mammalian limb development. *Dev. Cell* **19**, 819-830.
- Sparago, A., Cerrato, F., Vernucci, M., Ferrero, G. B., Silengo, M. C. and Riccio, A. (2004). Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome. *Nat. Genet.* **36**, 958-960.
- Spencer, R. J., Del Rosario, B. C., Pinter, S. F., Lessing, D., Sadreyev, R. I. and Lee, J. T. (2011). A boundary element between tsix and xist binds the chromatin insulator ctcf and contributes to initiation of x-chromosome inactivation. *Genetics* **189**, 441-454.
- Splinter, E., Heath, H., Kooren, J., Palstra, R. J., Klous, P., Grosveld, F., Galjart, N. and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* **20**, 2349-2354.
- Sun, F. L. and Elgin, S. C. (1999). Putting boundaries on silence. *Cell* **99**, 459-462.
- Szabo, P., Tang, S. H., Rentsendorj, A., Pfeifer, G. P. and Mann, J. R. (2000). Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function. *Curr. Biol.* **10**, 607-610.
- Udvardy, A., Maine, E. and Schedl, P. (1985). The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J. Mol. Biol.* **185**, 341-358.
- Vazquez, J. and Schedl, P. (2000). Deletion of an insulator element by the mutation facet-strawberry in Drosophila melanogaster. *Genetics* **155**, 1297-1311.
- Wan, L. B., Pan, H., Hannehalli, S., Cheng, Y., Ma, J., Fedoriw, A., Lobanekov, V., Latham, K. E., Schultz, R. M. and Bartolomei, M. S. (2008). Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development* **135**, 2729-2738.
- Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsumumi, S., Nagae, G., Ishihara, K., Mishiro, T. et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796-801.
- Weth, O., Weth, C., Bartkuhn, M., Leers, J., Uhle, F. and Renkawitz, R. (2010). Modular insulators: genome wide search for composite CTCF/thyroid hormone receptor binding sites. *PLoS ONE* **5**, e10119.
- Wood, A. M., Van Bortle, K., Ramos, E., Takenaka, N., Rohrbaugh, M., Jones, B. C., Jones, K. C. and Corces, V. G. (2011). Regulation of chromatin organization and inducible gene expression by a Drosophila insulator. *Mol. Cell* **44**, 29-38.
- Xiao, T., Wallace, J. and Felsenfeld, G. (2011). Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol. Cell. Biol.* **31**, 2174-2183.
- Yang, Y., Quitschke, W. W., Vostrov, A. A. and Brewer, G. J. (1999). CTCF is essential for up-regulating expression from the amyloid precursor protein promoter during differentiation of primary hippocampal neurons. *J. Neurochem.* **73**, 2286-2298.
- Yao, H., Brick, K., Evrard, Y., Xiao, T., Camerini-Otero, R. D. and Felsenfeld, G. (2010). Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev.* **24**, 2543-2555.
- Yu, W., Ginjala, V., Pant, V., Chernukhin, I., Whitehead, J., Docquier, F., Farrar, D., Tavosoidana, G., Mukhopadhyay, R., Kanduri, C. et al. (2004). Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat. Genet.* **36**, 1105-1110.
- Zhao, K., Hart, C. M. and Laemmli, U. K. (1995). Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* **81**, 879-889.
- Zhong, X. P. and Krangel, M. S. (1999). Enhancer-blocking activity within the DNase I hypersensitive site 2 to 6 region between the TCR alpha and Dad1 genes. *J. Immunol.* **163**, 295-300.
- Zhou, J. and Levine, M. (1999). A novel cis-regulatory element, the PTS, mediates an anti-insulator activity in the Drosophila embryo. *Cell* **99**, 567-575.